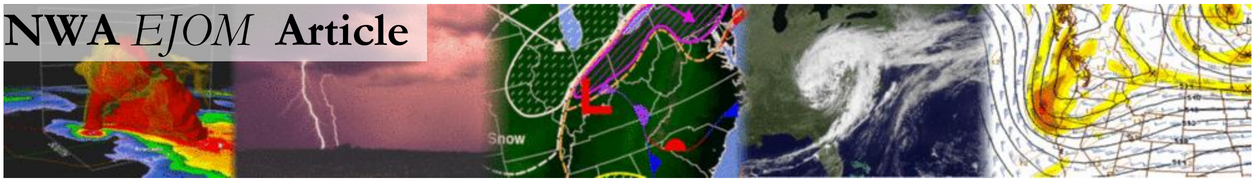
Blair, S. F., and J. W. Leighton, 2012: Creating high-resolution hail datasets using social media and post-storm ground surveys. *Electronic J. Operational Meteor.*, **13** (3), 32–45.



Creating High-Resolution Hail Datasets Using Social Media and Post-storm Ground Surveys

SCOTT F. BLAIR AND JARED W. LEIGHTON
National Weather Service, Topeka, Kansas

(Manuscript received 20 December 2011; in final form 9 April 2012)

ABSTRACT

Hail reports from *Storm Data* typically produce insufficient spatial and temporal resolution to determine the true hail-fall character of a storm. However, high-resolution hail databases are essential for meaningful hail studies utilizing radar base products and for the development and refinement of hail algorithms. Today, many broadcast and print media outlets provide supplemental social media forums where the public, armed with a wide array of digital cameras and mobile devices with photo and application software capabilities, can submit reports of severe weather. These technologies and social media web sites have the potential to serve as a substantial resource for additional meteorological observations. To illustrate the utility of reports from social media and post-storm ground surveys, hail information was gathered and analyzed from a notable hail event that occurred across the Wichita, Kansas, metropolitan area on 15 September 2010. A total of 464 hail size data points were obtained within a ~ 648 km^2 (250 mi^2) area, with 94% of the reports originating from social media and the hail survey. Additionally, social media and the post-storm ground survey identified eight hailstones that exceeded the diameter of the previous state record, with the largest diameter measured at 197 mm (7.75 in.). The reconstruction of the hail-fall character obtained from this dataset is among some of the highest spatial resolution hail datasets available to date, and has the additional benefit of photographic documentation for approximately 93% of the hail data points in the study.

1. Introduction

The National Weather Service (NWS) creates a verification database using severe weather reports gathered through warning verification efforts, ultimately published in *Storm Data*. These data frequently serve as the primary ‘ground-truth’ source of severe weather information in many convective weather and radar-based studies. Unfortunately, the imprecise and incomplete spatial and temporal resolution of the reports in *Storm Data* makes them difficult to use with confidence (Lenning et al. 1998; Witt et al. 1998; Marzban and Witt 2001; Blair et al. 2011). The traditional NWS verification practices are ultimately designed to efficiently verify severe weather warnings, not to satisfy scientific studies (Amburn and Wolf 1997). As a result

Corresponding author address: Scott F. Blair, National Weather Service, 1116 NE Strait Avenue, Topeka, KS 66616-1667
E-mail: scott.blair@noaa.gov

of the inherent low-resolution nature of *Storm Data*, there is a need for methods that improve the resolution in the severe weather database, especially for radar-based studies attempting to derive unique signatures from hail reports.

High-resolution hail databases are essential for meaningful hail research utilizing Weather Surveillance Radar-1988 Doppler (WSR-88D; Crum and Alberty 1993) legacy and dual-polarization base products, and for the development and refinement of hail algorithms. The representation of the hail-fall character of a storm increases with high-resolution hail databases. Studies utilizing these higher resolution reports are subject to minimal error compared to those only incorporating reports from *Storm Data* (Fig. 1). Currently, the only high-resolution hail datasets that adequately identify the hail-fall character in storms originate from the Severe Hazards Analysis and Verification Experiment (SHAVE; Ortega et al. 2009) and A Hail Spatial and Temporal Observing Network Effort (HailSTONE; http://hailstoneresearch.org/). While these efforts produce a wealth of information and insight into the character of hail storms, they unfortunately operate on a limited basis and over a limited domain area.

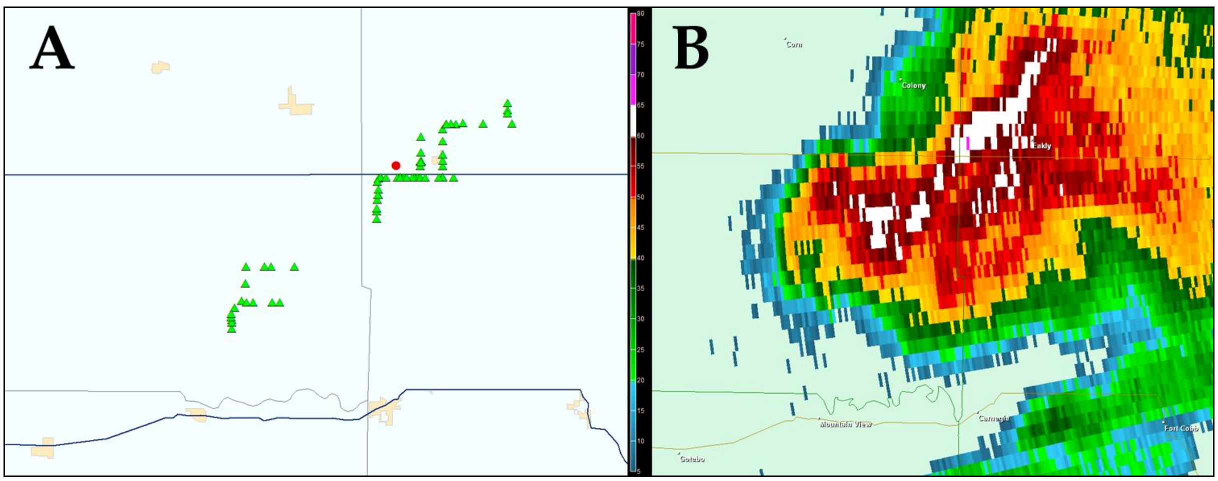


Figure 1. (a) An example of the spatial distribution of hail reports comparing high-resolution (HailSTONE, green triangle) versus low-resolution (*Storm Data*, red circle) hail datasets collected from 2000 UTC to 2030 UTC on 24 May 2011. (b) WSR-88D Oklahoma City (KTLX) 0.5^{o} base reflectivity image from 24 May 2011 at 2006 UTC used for reference.

In today's society, the public is armed with a wide array of mobile devices with photo and video capabilities that can easily share imagery on the Internet. Some of these technologies incorporate geotagging, which allows for location-specific information to be included within the imagery. Many media outlets encourage users to submit weather photos to their respective social media forums during weather events, resulting in an abundance of public weather-related information. In many cases, without this collection of social media information, the meteorological community would have insufficiently documented severe weather events and other weather phenomena (Hyvärinen and Saltikoff 2010). While the initial intentions of social media weather information are generally not aimed at scientific research, these observations in a quality-controlled environment have the potential to serve as a substantial resource in severe storm verification. This paper illustrates the utility of social media reports and additional verification efforts as resources in creating a high-resolution hail dataset from a notable hail event that occurred across the Wichita, Kansas, metropolitan area on 15 September 2010.

2. Hail data and methodology

a. Case example: 15 September 2010

On 15 September 2010, deep convection initiated over Reno County, Kansas, at approximately 2030 UTC. One long-lived supercell moved southeastward through portions of south-central Kansas over a 5-h period, impacting thousands of residents, including a large portion of the Wichita metropolitan area (Fig. 2). Most notably, the supercell produced an expansive swath of significant hail (diameter ≥ 51 mm; ≥ 2.00 in.) across both urban and rural areas. Additionally, the storm produced five EF-0 tornadoes in Sedgwick and Cowley Counties.

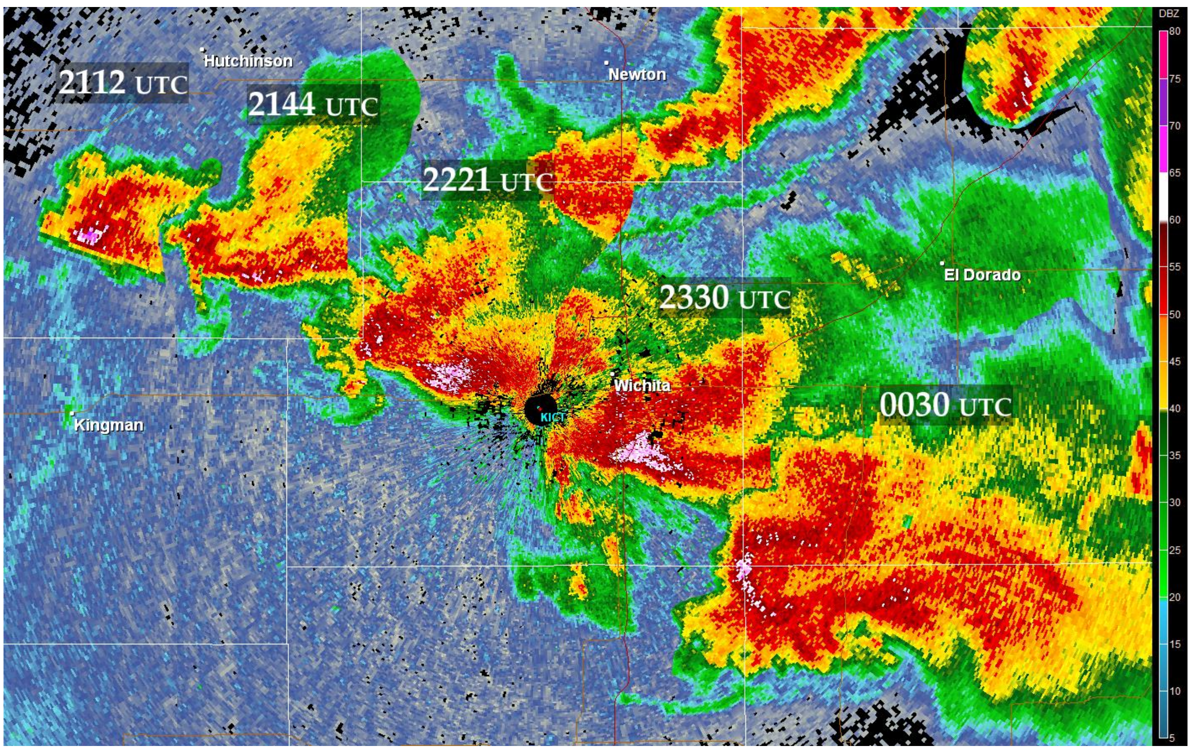


Figure 2. A time series composite of 0.5° base reflectivity from the Wichita, Kansas (KICT), WSR-88D from 2112 UTC 15 September 2010 to 0030 UTC 16 September 2010.

Residents identified a giant hailstone (hereafter referred to as the "Wichita stone") in the western suburbs of Wichita that had a maximum diameter of 197 mm (7.75 in.). The State Climate Extremes Committee (SCEC) certified the Wichita stone as the new state record diameter hailstone, exceeding the previous record of 144 mm (5.67 in.) that occurred in Coffeyville, Kansas, on 3 September 1970 (Fig. 3). The individuals who retrieved the Wichita stone took diameter measurements within 30 minutes of occurrence. In diameter, the Wichita stone ranks as the second largest verifiable United States hailstone to date. Table 1 lists the "Top 5" verifiable hail stones, derived from hail reports with valid photographic evidence or confirmed through a review process of the SCEC. An initial assessment from the hail event in Sedgwick County estimates $150 million in damages, with more than 35,000 claims turned in to insurance agencies (NCDC 2010).

Table 1. The "Top 5" largest verifiable hail stones by diameter in the United States. The rank, event location and date, maximum diameter and circumference, and a photo of the hailstone are listed below for each entry. An asterisk denotes measurements made after nontrivial melting occurred.

Rank	Location / Date	Diameter	Circumference	Photo
1	**Vivian, South Dakota** *23 July 2010*	8.00 in. (203 mm)	18.63 in. (473 mm)	
2	**Wichita, Kansas** *15 September 2010*	7.75 in. (197 mm)	15.50 in.* (394 mm)*	
3	**Aurora, Nebraska** *22 June 2003*	7.00 in. (178 mm)	18.75 in. (476 mm)	
4	**Dante, South Dakota** *21 August 2007*	6.88 in. (175 mm)	18.00 in. (457 mm)	
5	**Gotebo, Oklahoma** *23 May 2011*	6.00 in. (152 mm)	14.50 in.* (368 mm)*	
	Meadville, Missouri *24 May 2004*	6.00 in. (152 mm)	16.50 in. (419 mm)	
	Timpkin, Kansas *24 May 2011*	6.00 in. (152 mm)	N/A	

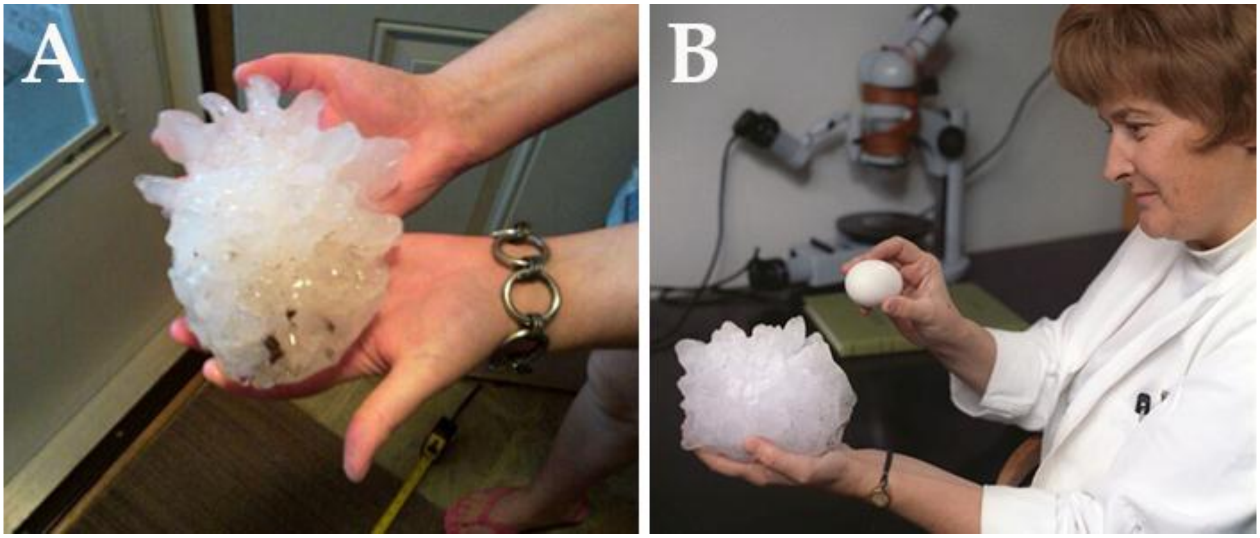


Figure 3. The (a) Wichita, Kansas, hailstone from 15 September 2010 with a maximum diameter of 197 mm (7.75 in.) compared to the (b) Coffeyville, Kansas, hailstone from 3 September 1970 with a maximum diameter of 144 mm (5.67 in.).

The data collected from the 15 September 2010 case incorporated all known available sources that contained hail information within the study domain, a 648 km^2 (250 mi^2) area centered over western Wichita, Kansas, that included the surrounding communities of Andale, Derby, Garden Plain, and Goddard. Hail data originated from three main sources: a) NWS local storm reports (LSRs), b) a post-storm ground survey, and c) several social media related web sites. The hail data were compiled in an Excel spreadsheet and imported into the Environmental Systems Research Institute's (ESRI) ArcMap by adding the latitude and longitude columns as x–y data. These data were then graphically depicted through an inverse distance weighted interpolation. The geographic information system (GIS) based maps of all hail reports collected provided tremendous insight into the hail-fall character of the storm.

b. NWS hail reports

To begin the construction of the hail database, NWS LSRs from 15 September 2010 were compiled. LSRs were utilized in place of *Storm Data* due to the higher number of available hail size data points. With this particular event, a single data entry in *Storm Data* incorporated many of the hail LSRs, providing only a start and end location of a singular swath of maximum hail size. Specific hail size information from the report location, including an improved temporal resolution, was available from the LSRs, making them the most attractive option in this case. Information that was collected from LSRs included the report time, measured or estimated maximum diameter hail size, and the location of the report.

c. Post-storm ground survey

A post-storm ground survey was conducted on 16 September 2010 after the NWS in Wichita, Kansas (NWS ICT) received several reports of giant hail (diameter ≥ 102 mm; ≥ 4.00 in.), including the Wichita stone. The objectives of the survey were to (1) assist NWS ICT with official measurements of the Wichita stone, (2) potentially identify any hail sizes larger than the Wichita stone, and (3) acquire additional hail reports within and in close proximity to the areas impacted by the largest hail in order to improve the report resolution for a separate radar-based study (Blair et al. 2011). A 24-km (15-mi.) long path of both rural and urban areas was surveyed from Garden Plain to the Wichita Mid-Continent Airport, including several subdivisions on the western side of Wichita. In the event that residents preserved hail in their freezers, the authors measured and recorded the diameter, circumference, and weight of each stone (Fig. 4). The authors also inquired whether residents made any hail reports to local authorities, media, or the NWS. Photographs were taken of each hailstone measured and the location was recorded.

Figure 4. An example of hailstones identified during the post-storm ground survey that were preserved by residents in the Wichita, Kansas, metropolitan area.

d. Social media

To supplement the reports obtained by the post-storm survey and NWS LSRs, hail information was gathered from eight different web-based social media outlets (Table 2). These web-based sources included local and national television stations, local print media, and dedicated social media sites that allowed video- and photo-sharing capabilities. In the context of this paper, social media are defined as interactive forums through which users could upload photos or video, share details, and view and comment on other's experiences from a particular event. The information contained within these social media pages provided the framework to document and record high spatial resolution hail information. The maximum diameter hail size was derived from each photo and video, using a 6.4-mm (0.25 in.) resolution. The majority of user-provided images compared the hailstone next to commonly-sized objects, tape measures, or rulers. For imagery that included no standard size comparison but included less traditional objects such as household or outdoor items, an estimated size was recorded. In the event that the hailstone was shown with no discernable object for size comparison, the report was removed from the database. Some inherent subjectivity was necessary for classifying specific maximum diameter hail sizes from amateur photos and videos, and additional uncertainty during the hail size classification stemmed from potential parallax errors. The size of the parallax error depended on the height and location of the hailstone relative to the camera lens. Individual cases were removed from the database when an approximate hail size could not be resolved due to considerable parallax errors. While the aforementioned methods of determining the maximum hail stone diameter allowed some uncertainty in the database with respect to the approximated diameters, the authors contend the data quality as a whole is likely superior to telephone-based reports of estimated hail sizes from the public due to the availability of photographic evidence and a consistent method for each case.

Table 2. Social media sources incorporated in the study that contained specific hail information from 15 September 2010.

Source	Web Address
Facebook (NWS Wichita)	www.facebook.com/US.NationalWeatherService.Wichita.gov/
iWitness Weather	iwitness.weather.com/
KAKE-TV	www.kake.com/
KSN-TV	www.ksn.com/
KWCH-TV	www.kwch.com/
Twitter	www.twitter.com/
The Wichita Eagle	www.kansas.com/
YouTube	www.youtube.com/

1) Broadcast media, print media, Facebook, and Twitter

Interactive photo galleries supplemented traditional broadcast and print media web pages in the Wichita metropolitan area. In these galleries, users were able to upload one or more photos of the event with the option to include a brief narrative caption (Fig. 5). There was also a section where individuals could leave comments about the images posted. Facebook and Twitter pages had a similar design of information-sharing. The authors searched for hail photos on the NWS Wichita Skywarn Facebook page, and on Twitter using the "hashtag" designation "#ksstorms." Each available photo was preserved, and data derived from the photos and narratives were recorded and entered into the database.

Photo captions frequently identified the location of the photographer, either by listing the intersection of two roadways or explicitly defining the address. In other cases it appeared that the photographers periodically approximated the location of the photograph using recognizable roadway intersections or landmarks closest to their respective location. Google Earth was utilized to locate the latitude and longitude coordinates of the specified hail location in the photo caption, and this information was quality controlled to determine if the provided location was reasonably accurate. In cases where no caption was listed or the address was unclear, the user's photo was removed from the database.

Figure 5. An example of hail photos uploaded by users on the KAKE social media photo forum from 15 September 2010.

2) YOUTUBE

The video sharing website YouTube (http://www.youtube.com) was interrogated for videos related to the 15 September 2010 hail storm. An initial search for "giant hail Wichita" was performed to identify videos taken and uploaded by residents and storm chasers, and videos of this event resulting from the search were then examined until all original videos were exhausted. Each YouTube account holder that uploaded a video from the hail storm was contacted with a request to provide the approximate time of the hail, the location where the video was taken, and the largest diameter of the hail in the video. The resulting information from this correspondence was utilized in the database.

3. Results and discussion

a. Reports, survey, and social media

1) NWS HAIL REPORTS

NWS LSRs accounted for 30 hail reports within the study domain. However, several of the report locations contained multiple reports of varying hail sizes as either stones became larger at a given area with time, or geographical-coordinated estimates placed reports in similar locations. Therefore, only 13 unique locations containing a maximum diameter hail size were produced from the LSRs.

2) POST-STORM GROUND SURVEY

The NWS post-storm ground survey established contact with 60 residents across the western sections of the Wichita metropolitan area. Approximately 80% of the surveyed residents preserved hailstones in their freezers. While some individuals placed hail in sealed plastic bags, the majority left stones exposed to the effects of sublimation. The residents indicated that the primary motivating factor of keeping the hailstones was for insurance purposes. Some individuals that preserved the stones also indicated that they did not collect the largest hail they observed. All stones saved by the residents had diameters ≥ 64 mm (≥ 2.50 in.), with a median size of 102 mm (4.00 in.). The high number of individuals found from the survey that saved hailstones yields some promise for future post-storm verification efforts in which residents may be able to provide accurate measurements of hail if explicitly prompted to do so.

Significant property damage was observed over the surveyed areas. The most impressive damage from hail impacts consisted of punctured automobile windows, large impact craters on the ground, and plywood roofs and wooden decks penetrated by giant hail.

A substantial find from the ground assessment was that no surveyed individuals contacted the NWS to report hail information, while approximately 2% reported a hail size to local authorities and 5% made reports to the media by phone or social media. These values illustrate the challenge of receiving real-time hail reports that provide an accurate assessment of the maximum hail size occurring in a storm, even in situations when highly urban landscapes are affected by significant hail. While the 5% of individuals reporting hail sizes to the media appears to be a low percentage, it emphasizes the contributing role that social media could play in providing additional verification information that would otherwise not be reported to a local NWS office.

3) Social media

Web-based social media sources accounted for 387 hail reports. Approximately 90% of the social media-based hail photos with a narrative caption originated from the three Wichita television broadcast station web pages. A hail size distribution of the social media reports revealed some insight into the reporting behavior of local residents (Fig. 6). The size distribution of hail reports obtained from photographic evidence deviates from traditional NWS verification practices that tend to bias hail sizes to commonly-sized objects or coins (Jewell and Brimelow 2009). The data also allow for some speculation that residents uploaded hail photos they felt were noteworthy to share, perhaps the largest hailstone within their proximity. It is interesting to note that only 3% of the uploaded photos contained hail smaller than the size of golf balls (44.5 mm; 1.75 in.) even though it is suspected that many urban locations, especially in the northeastern portion of the Wichita metropolitan area, received hail of this size. It is unknown whether the lack of reporting smaller hail sizes was a function of 1) a regional bias where large hail occurs more frequently and therefore smaller hail seems less notable, 2) a storm-specific bias where the coverage of significant hail in this storm was substantial and widespread, or 3) a seasonal bias where small hail during the convective season appears commonplace. More cases where social media plays a prominent role in hail verification will be required to assess the public's reporting behavior of hail sizes, and their perception of what constitutes a 'severe-sized' hailstone.

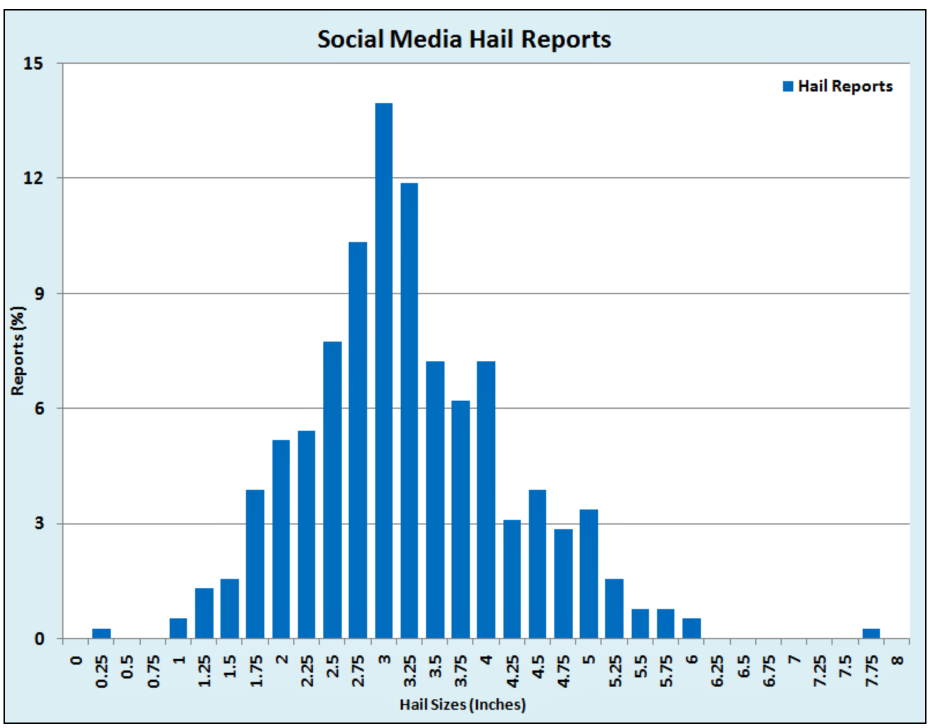


Figure 6. Normalized (%) hail size distribution (in.) of social media reports from 15 September 2010.

b. Recreating the hail-fall character

A total of 464 hail data points were collected from all available sources and incorporated into the database (Fig. 7). Social media information accounted for 83% of the reports, while 10% of the data points originated from the post-storm ground survey, and 7% of the points were from NWS LSRs. Approximately 93% of all the hail data points had accompanying photos of the maximum diameter hailstones. Hail reports with photographic documentation are superior to traditional reports, as the imagery helps mitigate uncertainty in the quality and validity of each hail data point. The tremendous number of available hail data points allowed for a high spatial resolution graphical reconstruction of the hail-fall character of the storm through time (Fig. 8). While an analysis of the hail-fall relative to radar-based signatures is beyond the scope of the paper, Figure 8 shows a decrease in the maximum hail size over time as the storm moved across the study domain, which would not have been apparent using only LSRs or *Storm Data*. This resolution of information is critical to accurately gauge hail-fall behavior and to help discriminate sensitive storm-scale changes that may be observed operationally by warning forecasters.

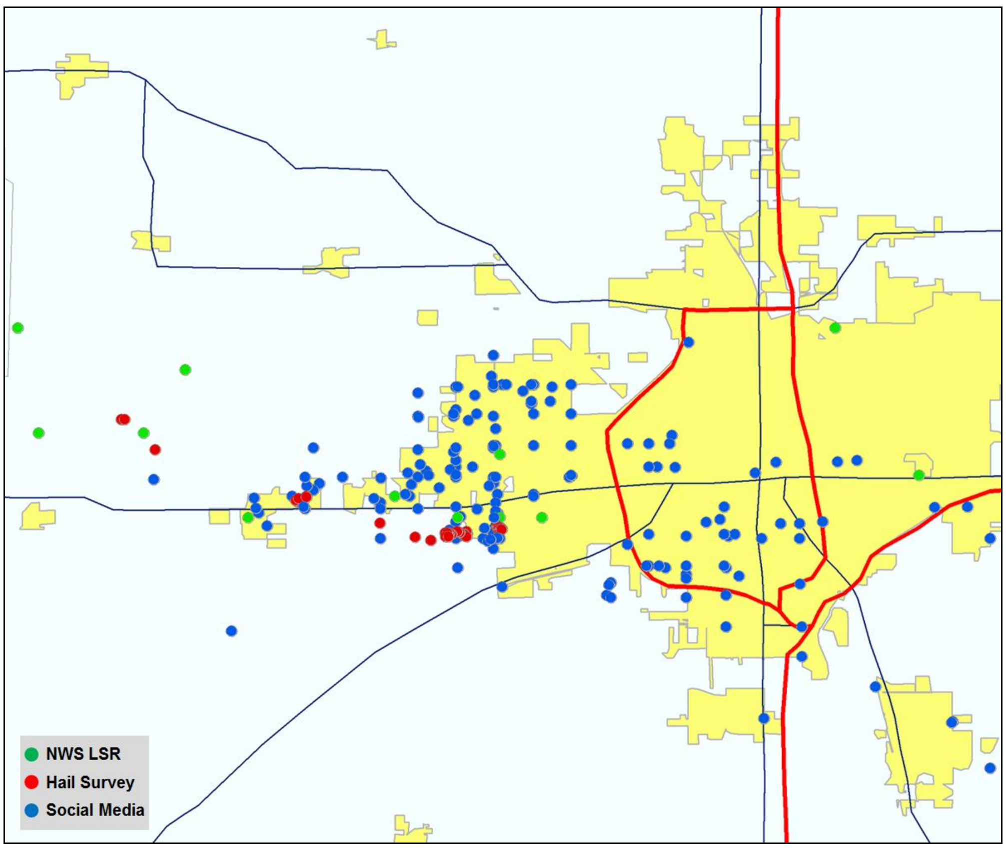


Figure 7. Locations of 464 hail data points from NWS LSRs (green dots), the post-storm hail survey (red dots), and social media (blue dots) utilized in the study. Urban areas (yellow) of the Wichita, Kansas, metropolitan area, primary Interstates (red lines), and state highways (blue lines) are shown for reference.

Eight hail stones, identified from the NWS post-storm ground survey and the search through social media-based web pages, exceeded the previous Kansas record diameter stone of 144 mm (5.67 in.). The residents that discovered the Wichita stone uploaded an image of the stone to one of the local television station's social media-based photo forums, and the station promptly shared the information with the NWS ICT during the ongoing severe weather event. It is worthy to note how important of a role such technologies served to document hailstones of this magnitude. While giant hail is a relatively rare phenomenon based on records in *Storm Data*, it has likely been underreported in the past (Blair et al. 2011). Table 1 reveals that the seven largest verifiable United States hailstones have occurred since 2003. Thus, it is reasonable to conclude more high-end hail events will continue to be documented as photo-sharing technologies become more widespread and utilized by the public and the meteorological community.

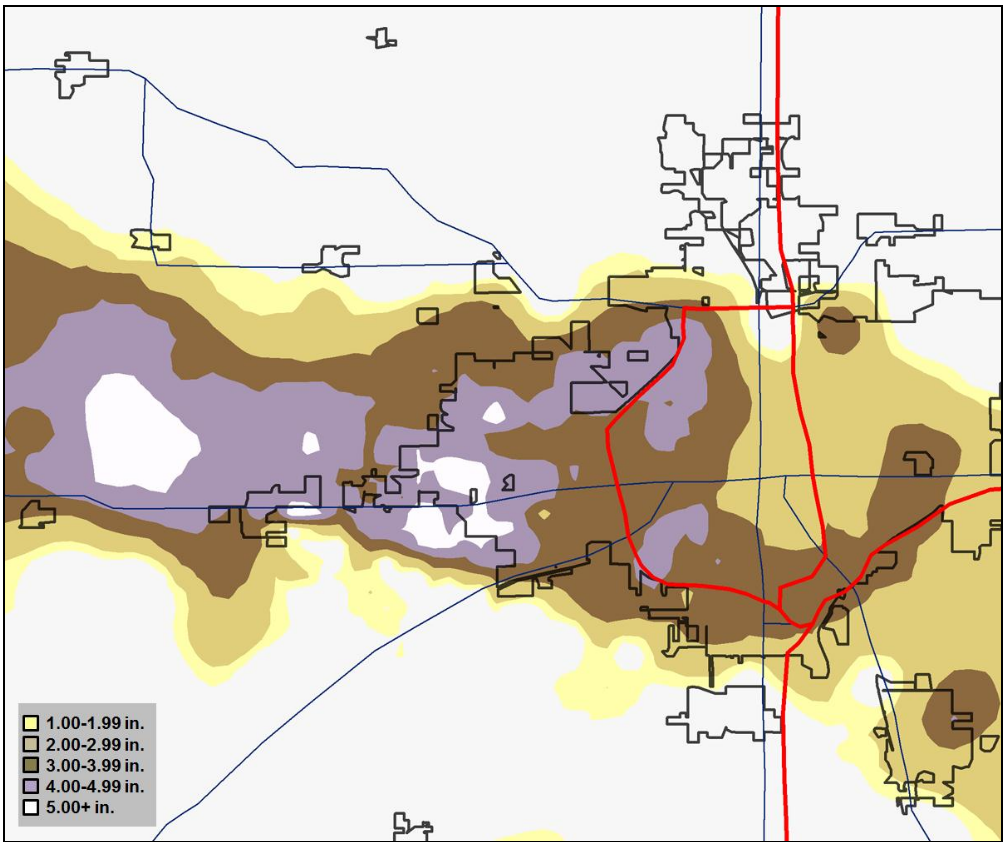


Figure 8. Inverse distance weighted interpolation from 464 hail data points. Color-filled contours correspond to one-inch increments of hail size. Urban areas (black outline) of the Wichita, Kansas, metropolitan area, primary Interstates (red lines), and state highways (blue lines) are shown for reference.

4. Summary

A high-resolution hail dataset was created from a prolific giant hail-producing supercell that impacted portions of the Wichita, Kansas, metropolitan area on 15 September 2010. The storm provided a unique opportunity to create very highly-spatially-resolved information within an urban and rural landscape, derived from supplemental hail data not available from *Storm Data*. The three main sources of hail data were NWS LSRs, a post-storm ground survey, and social media-related web sites. A total of 464 hail data points were collected and incorporated into the database, with 83% of the reports stemming from social media web sites. Approximately 93% of the data points had accompanying photos of the maximum diameter hailstones, mitigating some of the size errors and biases associated with typical public-based hail reports.

The NWS post-storm ground survey and social media data uncovered several useful findings. The survey found that approximately 80% of the sampled residents preserved hailstones in freezers, primarily for insurance purposes. Additionally, the survey showed the challenge of receiving hail reports from the public. None of the surveyed residents reported hail to the NWS, while approximately 7% contacted local authorities or the media. While this illustrates the challenge of receiving hail reports using traditional verification methods even when extraordinarily large hailstones occur in urban areas, the survey showed that social media outlets, especially those operated by local television stations, can be a significant source for obtaining additional severe weather reports. Lastly, the post-storm ground survey and searches through the Wichita-based social media web pages identified eight hail stones that exceeded the previous Kansas record hail size diameter.

Growing photo and video technologies and the usage of social media are allowing the operational and research communities to document meteorological events that would have gone unreported in the past. These data provide a meaningful contribution to the science, and are presumably available for most other notable weather phenomena. High-resolution hail data are critical to accurately represent the hail-fall behavior in a storm, in order to discriminate sensitive storm-scale changes and determine the utility of radar-based hail signatures. It is hoped that this type of supplemental hail data will be incorporated into *Storm Data* to improve the national hail climatology database, to support additional research efforts in hail size prediction utilizing dual-polarization radar products, and in the development and refinement of new hail-related algorithms.

Acknowledgements. The authors wish to recognize Jennifer Laflin (NWS Sioux Falls, SD), George Phillips (NWS Topeka, KS), Albert Pietrycha (NWS Goodland, KS), Derek Deroche (NWS Pleasant Hill, MO), Amos Magliocco (University of North Texas), Jeff Manion (NWS Central Region Headquarters), and two anonymous NWA reviewers for their role in improving the manuscript through technical editing. We would also like to thank Ken Cook and Chance Hayes (NWS Wichita, KS) for their assistance during the post-storm ground survey. Lastly, the authors thank the 2011 HailSTONE participants for their involvement in collecting data used in Fig. 1. The views expressed within this manuscript are those of the authors and do not necessarily represent those of the National Weather Service.

REFERENCES

Amburn, S. A. and P. L. Wolf, 1997: VIL density as a hail indicator. *Wea. Forecasting*, **12**, 473–478.

Blair, S. F., D. R. Deroche, J. M. Boustead, J. W. Leighton, B. L. Barjenbruch, and W. P. Gargan, 2011: A radar-based assessment of the detectability of giant hail. *Electronic J. Severe Storms Meteor*., **6** (7), 1–30.

Crum, T. D. and R. L. Alberty, 1993: The WSR-88D and the WSR-88D operational support facility. *Bull. Amer. Meteor. Soc.*, **74**, 1669–1687.

Hyvärinen, O. and E. Saltikoff, 2010: Social media as a source of meteorological observations. *Mon. Wea. Rev.*, **138**, 3175–3184.

Jewell, R. and J. Brimelow, 2009: Evaluation of Alberta Hail Growth Model Using Severe Hail Proximity Soundings from the United States. *Wea. Forecasting*, **24**, 1592–1609.

Lenning, E., H. E. Fuelberg, and A. I. Watson, 1998: An evaluation of WSR-88D severe hail algorithms along the northeastern Gulf Coast. *Wea. Forecasting*, **13**, 1029–1045.

Marzban, C. and A. Witt, 2001: A Bayesian neural network for severe-hail size prediction. *Wea. Forecasting*, **16**, 600–610.

NCDC, 2010: *Storm Data*. [Available from National Climatic Data Center, 151 Patton Ave., Asheville, NC 28801-5001.]

Ortega, K. L, T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. G. Kolodziej, and J. J. Gourley, 2009: The severe hazards analysis and verification experiment. *Bull. Amer. Meteor. Soc.*, **90**, 1519-1530.

Witt, A., M. D. Eilts, G. J. Stumpf, E. D. Mitchell, J. T. Johnson, and K. W. Thomas, 1998: Evaluating the performance of WSR-88D severe storm detection algorithms. *Wea. Forecasting*, **13**, 513–518.